Lan Li
1005814326

Exploring Dementia

1. Introduction

This project embarks on a meticulous exploration of MRI results from a longitudinal study focused on understanding the progression of brain volume changes in patients with and without dementia. This report will be based on the dataset called 'INF2178_A4_data.csv'.

The paper will be around two research questions:

1. Does the normalized whole brain volume (nWBV) change over time (between two visits), and is this change different between males and females?
2. How does nWBV vary across visits and does this vary by dementia status (Group: Nondemented, Demented, Converted)?
3. What is the appropriate sample size needed to detect a medium-to-large effect (effect size = 0.7) in nWBV between groups with 91% power and at a 5% significance level?

The questions can help us understand the factors which is affecting the whole brain volume of the patient.

2. Data Wrangling

The dataset totally includes **15 columns** and **294 rows**, but this report only investigates part of the columns. The columns (SES and MMSE) have missing values, thus there is a data cleaning needed. The cleaned data has 279 rows.

- Subject ID: A unique identifier assigned to each participant in the study to track individual data;
- Group: Dementia status, such as 'Nondemented', 'Demented', or 'Converted';
- Visit: Visit order of the patient (1 or 2);
- M/F: Gender of the patient (Male or Female);
- nWBV: Normalize Whole Brain Volume, the percentage of the intracranial cavity occupied by brain

3. Exploratory Data Analysis (EDA)

The interaction plot in **Figure** 1 shows that both males and females show a decrease in nWBV from Visit 1 to Visit 2. The male nWBV always seems lower than the one of female. There is no crossing of lines between Visits 1 and 2, suggesting there may not be a significant interaction effect between gender and visit on nWBV.

In Figure 2, all groups show a decline in nWBV over time, which could be associated with normal aging or disease progression. The converted group shows more variance than others.
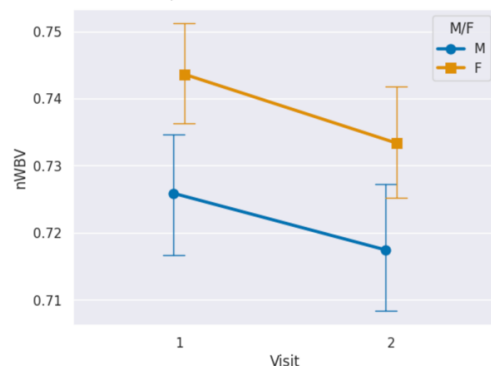

Visit	M/F	mean	std
1	F	0.74	0.04
1	M	0.73	0.04
2	F	0.73	0.04
2	M	0.72	0.04

Figure 1: Interaction Plot (Visit and Gender)

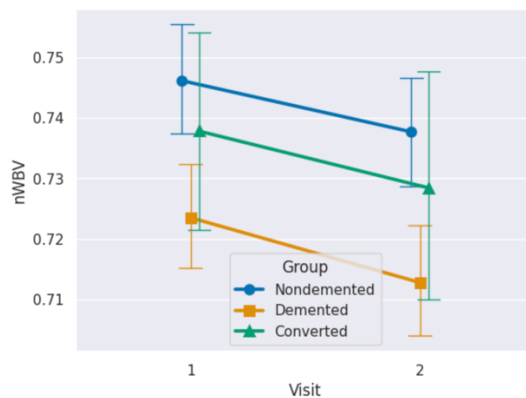| Visit	Group	mean	std
1	Converted	0.74	0.03
1	Demented	0.72	0.03
1	Nondemented	0.75	0.04
2	Converted	0.73	0.04
2	Demented	0.71	0.03
2	Nondemented	0.74	0.04

Figure 2: Interaction Plot (Visit and Group)

4. Mixed-Effect ANOVA

Research Question 1: Does the normalized whole brain volume (nWBV) change over time (between two visits), and is this change different between males and females?

The p-value (0.010) of M/F (Between-Subjects) means there is a significant effect of gender on nWBV, but it is relatively small in terms of effect size. The p-value of Visit (Within-Subjects) is less than 0.001 exhibiting there is a significant and strong effect of Visit on nWBV, indicating that nWBV changes significantly between the two visits. However, there is no significant interaction effect between gender and Visit (p=0.432), meaning that the change in nWBV from one visit to the next does not differ significantly between males and females.

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
M/F	0.019	1	135	0.019	6.902	0.010	0.049	nan
Visit	0.006	1	135	0.006	88.308	<0.001	0.395	1.000
Interaction	<0.001	1	135	<0.001	0.621	0.432	0.005	nan

Figure 3: Mixed ANOVA Summary (Visit and Gender)

Contrast	Visit	A	B	Paired	Parametric	T	dof	alternative	p-unc
Visit	-	1	2	TRUE	TRUE	9.410	136.000	two-sided	1.669E-16
M/F	-	F	M	FALSE	TRUE	2.615	116.517	two-sided	0.010
Visit * M/F	1	F	M	FALSE	TRUE	2.769	117.838	two-sided	0.007
Visit * M/F	2	F	M	FALSE	TRUE	2.409	116.234	two-sided	0.018

Figure 4: Post Hoc Test (Visit and Gender)

Research Question 2: How does nWBV vary across visits and does this vary by dementia status (Group: Nondemented, Demented, Converted)?

For Group (Between-Subjects), the p-value is 0.002, which is significant ($p < 0.05$), indicating there are statistically significant differences in nWBV between the different groups. For Visit (Within-Subjects), the p-value is less than 0.001, showing a highly significant effect of the visit on nWBV. The interaction effect has a p-value of 0.200, which is not significant ($p > 0.05$), indicating that the effect of the visit on nWBV does not significantly differ across groups.

Source	SS	DF1	DF2	MS	F	p-unc	np2	eps
Group	0.03	2	134	0.017	6.384	0.002	0.087	nan
Visit	0.01	1	134	0.006	89.376	<0.001	0.400	1
Interaction	0	2	134	<0.001	1.630	0.200	0.024	nan

Figure 5: Mixed ANOVA Summary (Visit and Group)

Contrast	Visit	A	B	Paired	Parametric	T	dof	alternative	p-unc
Visit	-	1	2	TRUE	TRUE	9.410	136.000	two-sided	0.000
Group	-	Converted	Demented	FALSE	TRUE	1.436	15.523	two-sided	0.171
Group	-	Converted	Nondemented	FALSE	TRUE	-0.648	15.832	two-sided	0.527
Group	-	Demented	Nondemented	FALSE	TRUE	-3.620	121.919	two-sided	0.000
Visit * Group	1	Converted	Demented	FALSE	TRUE	1.467	15.433	two-sided	0.162
Visit * Group	1	Converted	Nondemented	FALSE	TRUE	-0.465	15.995	two-sided	0.648
Visit * Group	1	Demented	Nondemented	FALSE	TRUE	-3.343	122.417	two-sided	0.001
Visit * Group	2	Converted	Demented	FALSE	TRUE	1.368	15.789	two-sided	0.190
Visit * Group	2	Converted	Nondemented	FALSE	TRUE	-0.811	15.615	two-sided	0.429
Visit * Group	2	Demented	Nondemented	FALSE	TRUE	-3.784	120.526	two-sided	0.000

Figure 6: Post Hoc Test (Visit and Group)

Assumption
Within the Shapiro-Wilk test, the large p-values (0.397 and 0.346 are typically larger than 0.05) suggests that there is no evidence to reject the null hypothesis of normality. For Levene's test. A high p-value (0.525) indicates that there is no evidence to reject the null hypothesis of equal variances across groups. Since these assumptions are met, it's appropriate to proceed with the ANOVA to examine the effects of Visit, M/F (or Group), and their interaction on nWBV.

Visit	W	pval	Normal
1	0.989894	0.397	True
2	0.988930	0.346	True

	W	pval	Normal
levene	0.404703	0.525	True

Figure 7: Assumption test of ANOVA

5. Power Analysis

Research Question 3: What is the appropriate sample size needed to detect a medium-to-large effect (effect size = 0.7) in nWBV between groups with 91% power and at a 5% significance level?

Figure 8 provides an overview of the appropriate sample size when power=0.91, alpha=0.05, and effect size=0.7. The sample size should be between 40 and 60. Through the calculation, the appropriate sample size should be at least 46 (accurately 45.451).
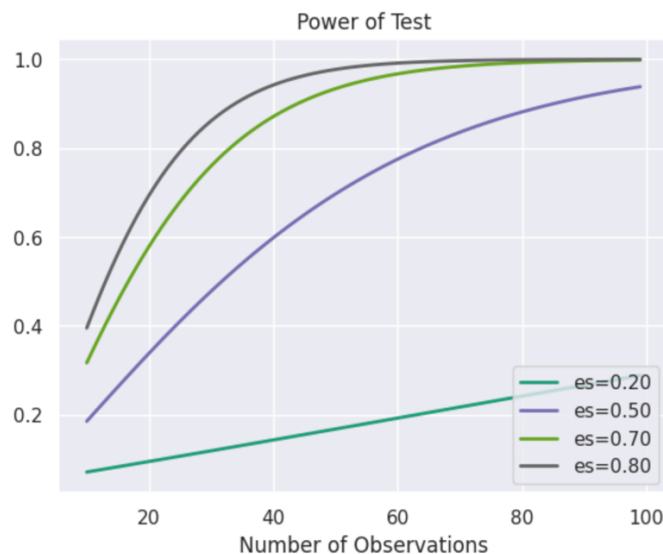

Figure 8: Power Analysis Plot